

NOV 07 2002

TECH CENTER 1600/2900



160C

11-12-a

RAW SEQUENCE LISTING

DATE: 10/31/2002

PATENT APPLICATION: US/09/765,555A

TIME: 17:33:34

Input Set : D:\27801-20014.txt

Output Set: N:\CRF4\10312002\I765555A.raw

ENTERED

3 <110> APPLICANT: The Scripps Research Institute
4 Barbas, Carlos
5 Stege, Justin
6 Guan, Xueni
7 Dalmia, Bipin
9 <120> TITLE OF INVENTION: Methods and compositions to modulate
10 expression in plants
12 <130> FILE REFERENCE: 27801-20014.20
14 <140> CURRENT APPLICATION NUMBER: 09/765,555A
15 <141> CURRENT FILING DATE: 2001-01-19
17 <150> PRIOR APPLICATION NUMBER: 09/620,897
18 <151> PRIOR FILING DATE: 2000-07-21
20 <150> PRIOR APPLICATION NUMBER: US 60/177,468
21 <151> PRIOR FILING DATE: 2000-01-21
23 <160> NUMBER OF SEQ ID NOS: 75
25 <170> SOFTWARE: FastSEQ for Windows Version 4.0
27 <210> SEQ ID NO: 1
28 <211> LENGTH: 532
29 <212> TYPE: DNA
30 <213> ORGANISM: Artificial Sequence
32 <220> FEATURE:
33 <223> OTHER INFORMATION: Promoter CsVMV
35 <400> SEQUENCE: 1
36 tctagaaact agcttcaga aggttaattat ccaagatgta gcatcaagaa tccaatgttt 60
37 acgggaaaaa ctatggaagt attatgtgag ctacagcaaga agcagatcaa tatgcggcac 120
38 atatgcaacc tatgttcaaa aatgaagaat gtacagatac aagatcctat actgccagaa 180
39 tacgaagaag aatacgtaga aattgaaaaa gaagaaccag gcgaagaaaa gaatcttgaa 240
40 gacgtaagca ctgacgacaa caatgaaaag aagaagataa ggtcggtgat tgtgaaagag 300
41 acatagagga cacatgtaag gtggaaaatg taagggcgga aagtaacctt atcacaaagg 360
42 aatcttatcc cccactactt atccttttat atttttccgt gtcatttttg cccttgagtt 420
43 ttcctatata aggaaccaag ttcggcattt gtgaaaacaa gaaaaaattt ggtgtaagct 480
44 attttctttg aagtactgag gatacaactt cagagaaatt tgtaagtttg ta 532
46 <210> SEQ ID NO: 2
47 <211> LENGTH: 18
48 <212> TYPE: DNA
49 <213> ORGANISM: Artificial Sequence
51 <220> FEATURE:
52 <223> OTHER INFORMATION: Zinc finger protein 2C7 binding site
54 <400> SEQUENCE: 2
55 gcgtgggagg cgtgggag 18
57 <210> SEQ ID NO: 3
58 <211> LENGTH: 51
59 <212> TYPE: DNA

RAW SEQUENCE LISTING

DATE: 10/31/2002

PATENT APPLICATION: US/09/765,555A

TIME: 17:33:34

Input Set : D:\27801-20014.txt

Output Set: N:\CRF4\10312002\I765555A.raw

```

60 <213> ORGANISM: Artificial Sequence
62 <220> FEATURE:
63 <223> OTHER INFORMATION: Promoter pc7rbTATA
65 <400> SEQUENCE: 3
66 cccgggtata taataagctt ggcattccgg tactgttggt aaagccacca t 51
68 <210> SEQ ID NO: 4
69 <211> LENGTH: 3121
70 <212> TYPE: DNA
71 <213> ORGANISM: Artificial Sequence
73 <220> FEATURE:
74 <223> OTHER INFORMATION: pND3008 coding region
76 <400> SEQUENCE: 4
77 agcgtgaccc ggtcgtgccc ctctctagag ataatgagca ttgcatgtct aagttataaa 60
78 aaattaccac atattttttt tgtcacactt gtttgaagtg cagtttatct atctttatac 120
79 atatatttaa actttactct acgaataata taatctatag tactacaata atatcagtgt 180
80 tttagagaat catataaatg aacagttaga catgggtctaa aggacaattg agtattttga 240
81 caacaggact ctacagtttt atcttttttag tgtgcatgtg ttctcctttt tttttgcaaa 300
82 tagcttcacc tatataatac ttcatccatt ttatttagtac atccatttag ggttttaggt 360
83 taatggtttt tatagactaa ttttttttagt acatctatct tattctatct tagcctctaa 420
84 attaagaaaa ctaaaactct atttttagtt ttttatttaa taatttagat ataaaataga 480
85 ataaaataaa gtgactaaaa attaaacaaa taccctttta gaaattaaaa aaactaagga 540
86 aacatttttc ttgttttcgag tagataatgc cagcctgtta aacgccgtcg acgagtcctaa 600
87 cggacaccaa ccagcgaacc agcagcgtcg cgtcgggcca agcgaagcag acggcacggc 660
88 atctctgtcg ctgcctctgg acccctctcg agagtccgc tccaccgttg gacttgctcc 720
89 gctgtcggca tccagaaatt gcgtggcgga gcggcagacg tgagccggca cggcaggcg 780
90 cctcctctc ctctcacggc acggcagcta cgggggattc ctttcccacc gctccttcgc 840
91 tttcccttcc tcgcccgcg taataaatag acaccccctc cacacccctc tttcccaacc 900
92 tcgtgtgttt cggagcgcac acacacacaa ccagatctcc cccaaatcca cccgtcggca 960
93 cctcgccttc aaggtacgcc gctcgtcctc cccccccccc cctctctacc ttctctagat 1020
94 cggcgttcgc gtcctatggt agggcccggt agttctactt ctgttcatgt ttgtgttaga 1080
95 tccgtgtttg tgtttagatcc gtgctgctag cgttcgtaca cggatgcgac ctgtacgtca 1140
96 gacacgttct gattgctaac ttgccagtgt ttctcttttg ggaatcctgg gatggctcta 1200
97 gccgttcgcg agacgggacg gatttcatga ttttttttgt ttcgttgcat aggggttggt 1260
98 ttgccctttt cctttatttc aatatatgcc gtgcacttgt ttgtcgggtc atcttttcat 1320
99 gctttttttt gtcttggttg tgatgatgtg gtctggttg gcggtcgttc tagatcggag 1380
100 tagaattctg tttcaaacta cctggtggat ttattaattt tggatctgta tgtgtgtgcc 1440
101 atacatattc atagttacga attgaagatg atggatggaa atatcgatct aggataggta 1500
102 tacatgttga tgcgggtttt actgatgcat atacagagat gctttttgtt cgttggttg 1560
103 tgatgatgtg gtgtggttg gcggtcgttc attcgttcta gatcggagta gaatactgtt 1620
104 tcaaactacc tgggtgtatt attaatattg gaactgtatg tgtgtgtcat acatcttcat 1680
105 agttacagat ttaagatgga tggaaatatc gatctaggat aggtatacat gttgatgtgg 1740
106 gttttactga tgcataata tgatggcata tgcagcatct attcatatgc tctaaccttg 1800
107 agtacctatc tattataata aacaagtatg ttttataatt attttgatct tgatatactt 1860
108 ggatgatggc atatgcagca gctatatgtg gattttttta gccctgcctt catacgtat 1920
109 ttatttgctt ggtactgtt cttttgtcga tgcacacctt gttgtttggt gttacttctg 1980
110 caggctcact ctagaggatc tatggcccag gcggccctcg agctccccta tgcttgccct 2040
111 gtcgagtcct gcgatcgccg cttttctaag tcggctgac tgaagcgcca tatcgcac 2100
112 cacacaggcc agaagccctt ccagtgtcga atatgcatgc gtaacttcag tcgtagtga 2160
113 caccttacca cccacatccg caccacacac ggcgagaagc cttttgcctg tgacatttgt 2220

```

RAW SEQUENCE LISTING

DATE: 10/31/2002

PATENT APPLICATION: US/09/765,555A

TIME: 17:33:34

Input Set : D:\27801-20014.txt

Output Set : N:\CRF4\10312002\I765555A.raw

114	gggaggaagt	ttgccaggag	tgatgaacgc	aagaggcata	ccaaaatcca	taccggtgag	2280
115	aagccctatg	cttgccctgt	cgagtcctgc	gacgcccgt	tttctaagtc	ggctgatctg	2340
116	aagcgccata	tccgcatcca	cacaggccag	aagcccttcc	agtgtcgaat	atgcatgcgt	2400
117	aacttcagtc	gtagtgaaca	ccttaccacc	cacatccgca	cccacacagg	cgagaagcct	2460
118	tttgccctgtg	acatttgttg	gaggaagtgt	gccaggagtg	atgaacgcaa	gaggcatacc	2520
119	aaaatccatt	taagacagaa	ggactctaga	actagtggcc	aggccggcca	ggctagcccg	2580
120	aaaaagaaac	gcaaagtgtg	gcgcgcccgc	gcgctggacg	atttcgatct	cgacatgctg	2640
121	ggttctgatg	ccctcgatga	ctttgacctg	gatagtgttg	gaagcgacgc	attggatgac	2700
122	tttgatctgg	acatgctcgg	ctccgatgct	ctggacgatt	tcgatctcga	tatgttaatt	2760
123	aactaccogt	acgacgttcc	ggactacgct	tcttgagaat	tcgcggccgc	gggcccgcgc	2820
124	ctagggagga	gctcaagatc	ccccgaatct	ccccgatcgt	tcaaacattt	ggcaataaag	2880
125	tttcttaaga	ttgaatcctg	ttgccggctc	tgcatgattt	atcatctaat	ttctgttgaa	2940
126	ttacgttaag	catgtaataa	ttacatgta	atgcatgacg	ttatttatga	gatgggtttt	3000
127	tatgattaga	gtcccgcgat	tatacattta	atacgcgata	gaaaacaaaa	tatagcgcgc	3060
128	aaactaggat	aaattatcgc	gcgcgggtgc	atctatgtta	ctagatccgg	gaattgggta	3120
129	c						3121
131	<210>	SEQ ID NO: 5					
132	<211>	LENGTH: 3069					
133	<212>	TYPE: DNA					
134	<213>	ORGANISM: Artificial Sequence					
136	<220>	FEATURE:					
137	<223>	OTHER INFORMATION: pND3018 coding region					
139	<400>	SEQUENCE: 5					
140	agcgtgaccc	ggcgtgccc	ctctctagag	ataatgagca	ttgcatgtct	aagttataaa	60
141	aaattaccac	atattttttt	tgtcacactt	gtttgaagtg	cagtttatct	atctttatac	120
142	atataattta	actttactct	acgaataata	taattctatg	tactacaata	atatcagtgt	180
143	tttagagaa	catataaatg	aacagttaga	catggtctaa	aggacaattg	agtattttga	240
144	caacaggact	ctacagtttt	atctttttag	tgtgcatgtg	ttctcctttt	tttttgcaaa	300
145	tagcttcacc	tatataatac	ttcatccatt	ttattagtac	atccatttag	ggtttagggg	360
146	taatggtttt	tatagactaa	tttttttagt	acatctatct	tattotattt	tagcctctaa	420
147	attaagaaaa	ctaaaactct	attttagttt	ttttatttaa	taatttagat	ataaaataga	480
148	ataaaataaa	gtgactaaaa	attaaacaaa	taccctttta	gaaattaaaa	aaactaagga	540
149	aacatttttc	ttgtttcgag	tagataatgc	cagcctgtta	aacgcgcgtc	acgagtctaa	600
150	cggacaccaa	ccagcgaacc	agcagcgtcg	cgtcgggcca	agcgaagcag	acggcacggc	660
151	atctctgtcg	ctgcctctgg	acccctctcg	agagttccgc	tccaccgttg	gacttgctcc	720
152	gctgtcggca	tccagaaatt	gcgtggcgga	gcggcagacg	tgagccggca	cggcaggcgg	780
153	cctcctcctc	ctctcacggc	acggcagcta	cgggggattc	ctttcccacc	gctccttcgc	840
154	tttccccttc	tcgcccgcgg	taataaatag	acaccccctc	cacacccctc	ttcccacaac	900
155	tcgtgttggt	cggagcgcac	acacacacaa	ccagatctcc	ccaaaatcca	cccgctggca	960
156	cctccgcttc	aagggtacgc	gctcgtcctc	cccccccccc	cctctctacc	ttctctagat	1020
157	cggcgttccg	gtccatgggt	agggcccggt	agttctactt	ctgttcatgt	ttgtgttaga	1080
158	tccgtgtttg	tgtttagatc	gtgctgctag	cgttcgtaca	cggatgcgac	ctgtacgtca	1140
159	gacacgttct	gattgctaac	ttgccagtgt	ttctcttttg	ggaatcctgg	gatggctcta	1200
160	gccgttccgc	agacgggatc	gatttcatga	ttttttttgt	ttcgttgcat	aggggttggt	1260
161	ttgccctttt	cctttatttc	aataatagcc	gtgcacttgt	ttgtcgggtc	atcttttcat	1320
162	gctttttttt	gtcttggttg	tgatgatgtg	gtctgggttg	gcggctcgtc	tagatcggag	1380
163	tagaattctg	tttcaaacta	cctgggtgat	ttattaattt	tggatctgta	tgtgtgtgcc	1440
164	atacatattc	atagttacga	attgaagatg	atggatggaa	atatcgatct	aggataggta	1500
165	tacatgttga	tgcggtgttt	actgatgcat	atacagagat	gctttttgtt	cgtttgtgtg	1560

RAW SEQUENCE LISTING

DATE: 10/31/2002

PATENT APPLICATION: US/09/765,555A

TIME: 17:33:34

Input Set : D:\27801-20014.txt

Output Set: N:\CRF4\10312002\I765555A.raw

```

166 tgatgatgtg gtgtggttgg gcggtcggtc attcggttcta gatcggagta gaatactggt 1620
167 tcaaactacc tgggtgtattt attaatTTTT gaactgtatg tgtgtgtcat acatcttcat 1680
168 agttacgagt ttaagatgga tggaaatatc gatctaggat aggtatacat gttgatgtgg 1740
169 gttttactga tgcataatac tgatggcata tgcagcatct attcatatgc tctaaccctg 1800
170 agtacctatc tattataata aacaagtatg ttttataatt attttgatct tgatatactt 1860
171 ggatgatggc atatgcagca gctatatgtg gatTTTTTT gccctgcctt catacgtat 1920
172 ttatttgctt ggtactgttt cttttgtcga tgctcaccct gttgtttggt gttacttctg 1980
173 caggtcgact cttagaggatc cactagttag ccatgggcta gcatggccgc tgcctgctgc 2040
174 atgaacatcc agatgctgct cgaagccgct gattatctgg aacgccggga gcgcgaagcc 2100
175 gagcacggct acgccagcat gctgccatat ccgaaaaaga aacgcaaggt ggcccaggcg 2160
176 gccctcgagc tccccatatg ttgccctgtc gagtcctgcg atcgccgctt ttctaagtcg 2220
177 gctgatctga agcgccatat ccgcatccac acaggccaga agcccttcca gtgtcgaata 2280
178 tgcatgcgta acttcagtcg tagtgaccac cttaccaccc acatccgcac ccacacaggc 2340
179 gagaagcctt ttgctgtga catttggtgg aggaagtttg ccaggagtga tgaacgcaag 2400
180 aggcatacca aaatccatac cggtgagaag ccctatgctt gccctgtcga gtctctgcg 2460
181 cgccgctttt ctaagtcggc tgatctgaag cgccatatcc gcattccacac aggccagaag 2520
182 cccttccagt gtcgaatatg catgcgtaac ttcagtgcga gtgaccacct taccaccac 2580
183 atccgcaccc acacaggcga gaagcctttt gcctgtgaca tttgtgggag gaagtttgcc 2640
184 aggagtgatg aacgcaagag gcataccaaa atccatttaa gacagaagga ctctagaact 2700
185 agtggccagg ccggccagta cccgtacgac gttccggact acgcttcttg aaagcttgg 2760
186 accgagctcg gatcccccgga atttccccga tcgttcaaac atttggcaat aaagtttctt 2820
187 aagattgaat cctgttgccg gtcttgcgat gattatcatc taatttctgt tgaattacgt 2880
188 taagcatgta ataattaaca tgtaatgcat gacgttattt atgagatggg tttttatgat 2940
189 tagagtcccg caattataca ttttaatacg gatagaaaac aaaatatagc gcgcaacta 3000
190 ggataaatta tcgcgcgcgg tgtcatctat gttactagat ccgggaattc cggaccggta 3060
191 ccagcggcc

```

193 <210> SEQ ID NO: 6

194 <211> LENGTH: 156

195 <212> TYPE: DNA

196 <213> ORGANISM: Artificial Sequence

198 <220> FEATURE:

199 <223> OTHER INFORMATION: 6X2C7 binding site

201 <400> SEQUENCE: 6

202 cgtgctagcg cgtgggcggc gtgggcgaac aagcgtgggc ggcggtggcg aacaagcgtg 60

203 ggcggcgtgg gcgactagt ctacgcgctg ggcggcgtgg gcgaacaagc gtgggcggcg 120

204 tgggcgaaca agcgtggcg gcgtggcgga ctagt 156

206 <210> SEQ ID NO: 7

207 <211> LENGTH: 18

208 <212> TYPE: DNA

209 <213> ORGANISM: Artificial Sequence

211 <220> FEATURE:

212 <223> OTHER INFORMATION: ZFPap3

214 <400> SEQUENCE: 7

215 gatggagttg aagaagta 18

217 <210> SEQ ID NO: 8

218 <211> LENGTH: 21

219 <212> TYPE: DNA

220 <213> ORGANISM: Artificial Sequence

222 <220> FEATURE:

RAW SEQUENCE LISTING

DATE: 10/31/2002

PATENT APPLICATION: US/09/765,555A

TIME: 17:33:34

Input Set : D:\27801-20014.txt

Output Set: N:\CRF4\10312002\I765555A.raw

```

223 <223> OTHER INFORMATION: ZFP from -85 to -65
225 <400> SEQUENCE: 8
226 gcctccttcc tcctctcact c 21
228 <210> SEQ ID NO: 9
229 <211> LENGTH: 18
230 <212> TYPE: DNA
231 <213> ORGANISM: Artificial Sequence
233 <220> FEATURE:
234 <223> OTHER INFORMATION: ZFPm1 from -68 to -85
236 <400> SEQUENCE: 9
237 tgagaggagg aaggaggc 18
239 <210> SEQ ID NO: 10
240 <211> LENGTH: 18
241 <212> TYPE: DNA
242 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: ZFPm2 from -65 to -82
247 <400> SEQUENCE: 10
248 gagtgagagg aggaagga 18
250 <210> SEQ ID NO: 11
251 <211> LENGTH: 24
252 <212> TYPE: DNA
253 <213> ORGANISM: Artificial Sequence
255 <220> FEATURE:
256 <223> OTHER INFORMATION: ZFP from 294 to 317
258 <400> SEQUENCE: 11
259 gccaaactact acggctccct cacc 24
261 <210> SEQ ID NO: 12
262 <211> LENGTH: 18
263 <212> TYPE: DNA
264 <213> ORGANISM: Artificial Sequence
266 <220> FEATURE:
267 <223> OTHER INFORMATION: ZFPm3 from 311 to 294
269 <400> SEQUENCE: 12
270 ggagccgtag tagttggc 18
272 <210> SEQ ID NO: 13
273 <211> LENGTH: 18
274 <212> TYPE: DNA
275 <213> ORGANISM: Artificial Sequence
277 <220> FEATURE:
278 <223> OTHER INFORMATION: ZFPm4 from 317 to 300
280 <400> SEQUENCE: 13
281 ggtgaggagg ccgtagta 18
283 <210> SEQ ID NO: 14
284 <211> LENGTH: 3300
285 <212> TYPE: DNA
286 <213> ORGANISM: Artificial Sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/765,555A

DATE: 10/31/2002

TIME: 17:33:35

Input Set : D:\27801-20014.txt

Output Set: N:\CRF4\10312002\I765555A.raw

L:1103 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:52